Environment and Planning B: Planning and Design 2012, volume 39, pages 326 – 343

doi:10.1068/b37130

Investigating the implications of using alternative GIS-based techniques to measure accessibility to green space

Gary Higgs, Richard Fry, Mitchel Langford
GIS Research Centre, Wales Institute of Social and Economic Research, Data and Methods, Faculty of Advanced Technology, University of Glamorgan, Pontypridd CF37 1DL, Wales;
e-mail: ghiggs@glam.ac.uk, rfry@glam.ac.uk, mlangfor@glam.ac.uk
Received 14 September 2010; in revised form 22 March 2011; published online 19 January 2012

Abstract. A large body of research has examined relationships between accessibility to green space and a variety of health outcomes with many researchers finding benefits in terms of levels of physical activity and relationships with levels of obesity, mental health, and other health conditions. Such studies often use spatial analytical techniques to examine relationships between distance to such spaces and health data collated at an individual survey respondent's home address or, more commonly, derived from area-based census measures summarised at a centroid. Generally, such measures are becoming more sophisticated and have moved on from the use of straightforward Euclidean-based measures to those based on network distance. However, few studies tend to use a combination of approaches or seek to establish the implications of incorporating alternative measures of accessibility on potential relationships. Using a database of green spaces (and associated attributes) and a detailed network dataset for the city of Cardiff, Wales, we examine the sensitivity of findings to the ways in which different metrics are calculated. This is illustrated by examining the variations in association between such metrics and a census-based deprivation index widely used in health studies to measure socioeconomic conditions. Our findings demonstrate that not only will the distances to green spaces vary according to the methodologies adopted but that any study that aims to investigate relationships with attributes of the *nearest* green space should acknowledge that matches may vary widely according to the techniques used. We conclude by warning against the use of inappropriate methodologies in examining access to green space which may directly influence directions (and levels) of association and hence may limit their relevance in wider geographical contexts.

Keywords: green space, accessibility, GIS, spatial data, representation, network distances

1 Introduction

Studies of health inequalities and disparities within areas are increasingly concerned with the potential for health-promoting resources to impact upon causative behaviour (Humpel et al, 2002). It is within this context that the health and well-being benefits of access to green space have received growing recognition. Positive benefits for both mental and physical health are acknowledged through factors such as the promotion of physical activity and stress reduction (de Vries et al, 2003; Mitchell and Popham, 2008). Although the exact pathways are contested, access to green space may be particularly beneficial for children and lower socioeconomic groups for conditions such as anxiety disorder and depression (Maas et al, 2009). In previous studies such benefits are primarily determined by the quantity and characteristics of parks or recreation spaces and their proximity to potential users (Kaczynski and Henderson, 2007). Distance-based measures are often used to examine associations between green space and health outcomes controlling for potential confounders in any relationship. Such studies have tended to demonstrate that usage of green space declines with increased distance and time, with some noting a cut-off point of around 300 – 400 m as being a significant threshold (Giles-Corti et al, 2005; Nielsen and Hansen, 2007). Van Herzele and Wiedemann (2003, page 111) posit that "distance or walking time from the home has appeared to be the single most important precondition for use of green space."

Other studies have found that the likelihood of using parks for physical activity also relates to the features and range of facilities they offer, and that these factors are more important than size or accessibility in influencing levels of use (Kaczynski et al, 2008).

Geographical information systems (GIS) offer the potential to generate more accurate estimates of accessibility and a number of alternative methods of measuring distance are available. Forsyth (2000) drew attention to the inconsistencies in definitions, as well as gaps in available data, which can hinder the use of GIS for measuring access to public space. However, few studies have explored the implications of using alternative methods to calculate distance to green space for computed levels of association with socioeconomic or health measures. In this paper we examine the sensitivity of findings to the ways in which different distance metrics are calculated, using a green spaces database and detailed network dataset for the city of Cardiff, Wales. Specifically, we explore relationships between accessibility measures and the various ways in which spatial data can be represented and distances calculated within a GIS using the example of access to green space. Previous research such as that of Sander et al (2010) has drawn attention to the ways in which alternative methods of calculating the distance to open spaces significantly influence actual distances measured (in relation to the incorporation of distance variables in hedonic pricing models) by comparing Euclidean, road-network distances and raster-based cost-weighted distances. In reality such models may be even more complex if we consider the methods through which supply-side features (ie, green spaces in this instance) are represented, their location in relation to the existing road network and the ways in which distance is calculated using GIS-based network models. In this paper we build on that study to examine the ways in which a combination of different approaches to measuring distance may interact with how spatial features are actually represented in a GIS to influence the directions (and levels) of association in wider studies of the (potential) health-promoting benefits of green spaces.

The rest of this paper is structured as follows; in the next section we briefly review previous approaches to measuring access to green space focusing primarily on the health geography literature (whilst acknowledging increasing research efforts in the field of environmental (in)justice and access to green space). We highlight some of the key methodological issues that arise from such research before describing in section 3 the methodologies we have adopted to address some of these concerns. We next present our findings from a GIS and statistical-based analysis of access to green space and use a real-world example to illustrate the complexity surrounding measuring accessibility that has tended to be overlooked in previous research efforts. First this involves an examination of the relationships between the different metrics and, secondly, a comparison of the levels of association between such metrics and a census-based deprivation index that is widely used in health studies in the UK to measure socioeconomic conditions. In the final sections we highlight the importance of acknowledging different ways of measuring access to green space and call for researchers in this area to incorporate alternative measures when examining potential associations with socioeconomic or health data.

2 Spatial approaches to measuring access to green spaces

2.1 Previous studies

Spatial approaches have been widely used to examine access to green spaces in urban areas, often in the context of wider notions of environmental justice and the equity of provision. For example, in a national study in the USA Powell et al (2004) used outdoor observational data on community-level physical-activity-related settings, coupled with census data, to examine green space provision in relation to race and household income. Findings suggest that areas with higher poverty rates were significantly associated with reduced availability of green spaces, parks, and public sports areas, while areas with higher household income had a greater frequency of such amenities. Ethnic minority neighbourhoods, such as those with predominately African-American populations, were also associated with fewer areas of green spaces, parks, and sports areas. These findings contrast with other studies (eg, Jones et al, 2009; Macintyre et al, 2008) who found either no relationship between the provision of spaces and facilities with deprivation or ethnic minority neighbourhoods, or opposite trends. Barbosa et al (2007), for example, used GIS to measure distance along transport networks to public green spaces for households within Sheffield, UK. They report a general lack of accessible green space for Sheffield residents, but those with highest levels of potential access were the least affluent and elderly. Boone et al (2009) found that whilst African Americans in Baltimore had greater access to parks within a 400 m walking distance, the acreage of such areas available to whites within walking distance is greater, suggesting other factors need to be taken into account in such studies. However, few studies incorporate measures of utilisation of green space, and any comparison of findings is made more problematic by different definition of facilities and spaces used and the spatial definitions and methodologies adopted.

GIS have been widely used to examine the spatial pattern of accessible natural green space and to identify those areas currently lacking provision. Often these measures are used to investigate a possible ecological association between availability and accessibility of green space and health outcomes and have tended to use census tracts as the supporting spatial unit. Whilst this 'container-based' approach forms the basis for a number of ecological studies (eg Abercrombie et al, 2008; Mitchell and Popham, 2007) they have acknowledged limitations such as assuming people do not travel outside their administrative area to access green space (Talen and Anselin, 1998). In recognising these limitations, GIS researchers have more recently addressed such concerns in two main ways; (i) by buffering open spaces and finding population characteristics within buffered areas, or (ii) by using network analysis to take into account actual transport routes with assumed walking speeds. Often these are then integrated with nonspatial factors to investigate whether distributions are equitable. Direct (ie, buffer) distances almost always underestimate actual distances since individuals must travel along predefined rights of way to reach green spaces (Nicholls, 2001). In practice there is often a trade-off between the availability of appropriate datasets and software, and the aims/objectives and resources of the particular project, and a number of studies have used different neighbourhood measures to examine potential relationships with activity levels and health measures.

2.2 Methodological issues and concerns

A common approach in the use of GIS for measuring accessibility to green space is to calculate 'coverage' measures which report the number of services within a Euclidean distance of points of origin for those people potentially using such services. These tend to be represented as census tract population-weighted centroids or individual residences. Alternatively, GIS may be used to derive 'proximity' measures based on the population within specified drive-times of services. In either case there are three fundamental elements of interest: (1) an origin point, representing the geographical location of the population potentially seeking to access green space; (2) a destination point, representing the geographical location of the green space; and (3) a distance measurement taken between these two points. The precise formulation or identification of each of these three elements has the potential to raise issues and concerns.

Firstly, the way in which a population point of origin is estimated can be postulated to impact on accessibility measures. In studies of respondent surveys the postcode (zipcode) of an individual's home address can be used to compute actual distance to an entrance or another access point (or centroid) of the nearest green space (eg, Coombes et al, 2010) in order to, for example, compare perceived versus objective measures of accessibility. However, in the absence of individual household-level population data, the majority of studies to date tend to approximate points of origin using geometric or population-weighted centroids of census tracts as proxy locations of population demand for services. The limitations of this approach have been explored previously but with different outcomes; Hewko et al (2002) investigated the importance of 'aggregation-error effects' on a minimum-distance measure of neighbourhood accessibility and found that such effects vary according to the service or amenity under consideration. Jones et al (2010) found relatively small differences in the impact of geographic placement between resident addresses and zip-code centroids relative to the distance measurement technique employed. More research is needed to investigate whether such factors (ie, population representation) also impact on accessibility measures developed in green space research and this forms the basis of our ongoing research and as such is not included as part of this particular study.

Secondly, the manner in which green space is represented in a GIS may also impact on derived accessibility scores. A question such research poses relates to whether proximity should be measured to park entrances (eg, Barbosa et al, 2007), to other known access points (eg Jones et al, 2009), to the nearest green space boundary (eg, Kessel et al, 2009), or by Euclidean distance from residential addresses to park centroids (eg, Kaczynski et al, 2008). Previous studies have suggested that, unlike the traditional use of park centroids to measure access, using multiple entrance points enables a park's shape to be accounted for (Nicholls, 2001; Nicholls and Shafer, 2001). Comber et al (2008), for example, measure distance between such access points and UK output area (OA) centroids. However, more research is needed to investigate whether distance measures based on green space centroids are sufficiently robust to represent variations in accessibility and to examine the impacts of alternative measures in this context. Furthermore, the continued use of centroids within such studies can be contested given that more sophisticated techniques are now widely available within GIS packages. However, the extra effort involved in creating the necessary datasets (such as access points) often means researchers are limited to using centroids in their analysis. To date few studies have compared the implications of using different spatial representations in green space research—which therefore forms one of the research questions addressed in this paper.

Finally, we need to consider the measurement technique used to establish the distance between the chosen origin and destination points—particularly whether Euclidean distance or network distance (or time) is utilised. Some research has been conducted in the health arena, for example, on the efficacy of using network distance accessibility measures over straight-line distance, in both the UK (Jordan et al, 2004), and the US (Phibbs and Luft, 1995). A study of alternative approaches to examining access to a range of health services in Montreal, Canada (Apparicio et al, 2008) found straight-line distances correlate with more advanced network-based measures, but not uniformly across the metropolitan area. Similar effects were noted in respect to the way population demand was represented (eg, population-weighted versus geometric centroids). Little research has been conducted to compare such impacts on studies of access to green space to see if the 'suburban effect' found in the Montreal studies is replicated in other health studies. With regard to green space studies, network distance rather than drive time is postulated to be the most appropriate measure given that most people walk to their nearest green space.

Few studies use a combination of distance measurements or compare the implications of using one approach over another—and this forms another research question addressed in this paper. The availability of a rich dataset of green space boundaries and access points (which include park entrances) for the city of Cardiff as well as a detailed digital road network enabled us to test the sensitivity of accessibility measures to different parameters in our distance calculations. The methodological approach adopted is described in the next section.

3 Methodology

The broad approach taken in this study involves measuring distance to green space from population points of origin via a number of alternative methods and determining the identity of the nearest green space object for these different scenarios. Our aim is to evaluate the degree to which outcomes are influenced by the details of the approach adopted. As stated above there are three crucial elements of concern: (i) the object or geographical location acting as an origin point *from* which a distance is measured; (ii) the object or geographical location acting as a destination point *to* which a distance is measured; and (iii) the nature of the distance measurement taken between these two points.

3.1 Data sources

For the object or geographical location that defines the point of origin for green space distance measures we utilise population-weighted centroids for the 2001 UK Census of Population OAs, as supplied by the Office for National Statistics (http://www.statistics.gov.uk) OAs are the most detailed data aggregation units available within the UK census and contain around 80 – 150 households. Although alternative datasets are available from other sources that can offer an even finer disaggregation of population distribution, as we note above, the possible impacts of alternative population representations will be considered in future work. For the polygons that represent a target destination for the distance measured we utilise a database of accessible natural green spaces (totalling just under 600 polygons, in March 2008) provided by Cardiff City Council. This dataset includes sites in all ownerships and not just those of the local authority (see figure 1). For each space over 0.25 ha we have information on the name of the main location, its area in hectares, the category of the space (eg, amenity, informal, woodland, Site of Special Scientific Interest, Site of Importance for Nature Conservation, Local Nature Reserve) and additional information (eg, semiimproved calcareous grassland, seminatural woodland, mixed amenity). Full details of the sources of baseline data are documented by Cardiff County Council (2008) and include aerial photographs, site visits, and detailed Ordnance Survey data. The dataset excludes sites that are perceived to be 'inaccessible' or 'unnatural' such as children's playgrounds, locked school fields and yards, private tennis courts, and allotments, but includes golf courses which may have public rights of way or other access routes.

Whilst these green space polygons define the objects for which we seek to report a distance, there are a number of options associated with each green space that define the specific geographical location utilised during the measurement process. Firstly, we have *access points* created by identifying locations where paths indicated on detailed Ordnance Survey maps and aerial photography enter a natural green space, where public rights of way enter a site, or through local knowledge and site visits [see Cardiff County Council (2008) for further details]. Such access points will include the 'official' park entrance(s) to green spaces but are not necessarily confined to such entrances. Within our dataset a total of almost 2800 access points are thus defined (see figure 1).

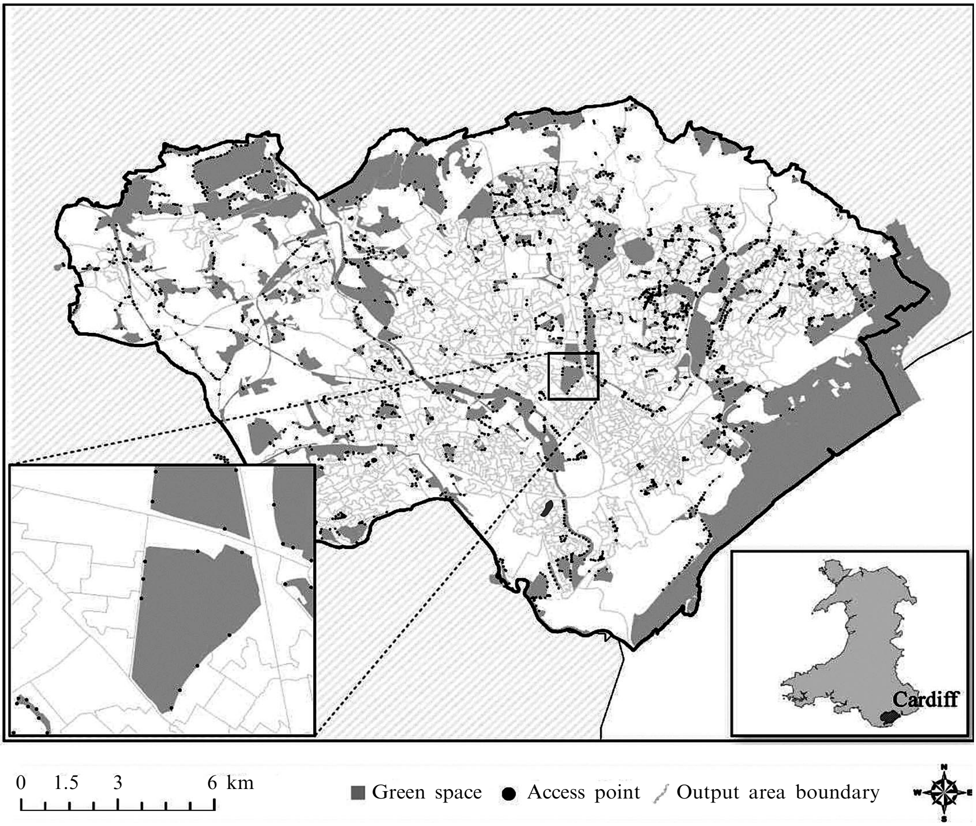


Figure 1. Distribution of green spaces in Cardiff.

Secondly, for each green space polygon we compute a geometric centroid (which, if necessary, is constrained to lie within the boundaries of the polygon). Boone et al (2009) suggest centroids of green space polygons can represent the destination points for small parks, but that otherwise parks will most probably have multiple entry points. Boone et al (2009, page 772) also suggest that "for larger parks, any point along the perimeter can arguably serve as the destination" and in many instances green spaces are accessible along most of their boundary. The third option, therefore, is based on the entire green space polygon boundary. Since we require a specific target point in order to measure a distance we compute the nearest (Euclidean distance) point on any green space boundary to each population-weighted centroid. Each green space has a unique identifier which permits the identity of the nearest polygon to be accessed and compared for the different metrics in the analysis that follows.

Finally, with respect to the nature of the measurement taken between OA centroids and green space polygons, ArcGIS™ 9.3 was used to calculate both Euclidean and network distance. Network distance was calculated in the Network Analyst extension software using the most detailed UK network dataset currently available (Ordnance Survey's Integrated Transport Network layer). No attempt was made at this stage to differentiate between street classes, or to incorporate attributes of these links or crossing points in the analysis, but clearly such information could be added to provide more 'realistic' estimates of routes taken to avoid major routes or barriers. As Apparcio et al (2008, page 7) suggest, "shortest network distance is useful for evaluating the path between two points as if taken on foot; consequently, it is frequently used in studies on the accessibility of 'proximal' services and facilities." Given our assumption that people walk to their nearest green space, the 'shortest distance' option was specified within Network Analyst rather than accepting the default drive-time option. Similarly, we deactivated any network parameters concerning one-way streets and turning restrictions at junctions. Since we do not have data for green space beyond the city of Cardiff we acknowledge there may be a small 'border effect' in our results.

3.2 GIS analysis

On the basis of our review of previous approaches to the measurement of access to green space, initially six methods for calculating proximity were identified from the literature (three used Euclidean distance and three network distance; each to the nearest green space centroid, boundary point, or access point, respectively)—all of which can be calculated from the centroid of an OA (see figure 2). It is important to note that in order to perform network distance analysis all origin and destination points must lie directly on the network dataset; this is typically not the case for either the OA centroids or any of the alternative green space destination points. To facilitate computation Network Analyst automatically moves all points that are used in an analysis to the nearest orthogonal position on the network. Typically, therefore, researchers use a tolerance value by using offsets to locate OA centroids and green space areas on the road network when calculating road network distances (Sander et al, 2010). Thus each OA centroid will typically be shifted by some distance, shown as X_1 in figure 2, to meet this requirement. Likewise, a green space centroid, nearest boundary point, or nearest access point will be shifted by some distance shown as X_3 in figure 2. The origin–destination matrix returned by Network Analyst reports only the network length between the shifted positions, or the distance shown as X_2 in figure 2. We suspect that most often it is this network distance which is used in subsequent investigations—because either the analyst is unaware that these shifting operations have occurred, or they do not wish to concern themselves with computing the distances X_1 and X_3, or they do not expect these components to have any significant impact upon results. However, some proximity studies have computed the 'full' network length consisting of the distance from the origin point to the transport network, the distance through the network itself, and the distance from the network to the destination point (ie, $X_1 + X_2 + X_3$). In this study we record both 'network path only' distance (ie, X_2) and 'full network path' distances (ie, $X_1 + X_2 + X_3$). Thus we compute nine distance measures in total for each OA centroid, with scenarios 4, 5, and 6 being repeated for each of these two approaches.

Scenario 1: Euclidean distance to the nearest centroid of any green space.
Scenario 2: Euclidean distance to the nearest boundary point of any green space.
Scenario 3: Euclidean distance to the nearest access point of any green space.
Scenario 4: Network distance to the nearest centroid of any green space ('network path only' and 'full network path').
Scenario 5: Network distance to the nearest boundary point of any green space ('network path only' and 'full network path').
Scenario 6: Network distance to the nearest access point of any green space ('network path only' and 'full network path').

The movement of the origin and destination points such that they lie on the network during network-based distance calculations can give rise to a number of possible complexities and side effects, some of which are illustrated in figure 3. In the first schematic [figure 3(a)] the green space labelled B is identified as being the closest based on Euclidean distance to a centroid (ie, $M < N$), but once the green space and OA centroids have been repositioned to lie on the network, green space A becomes the

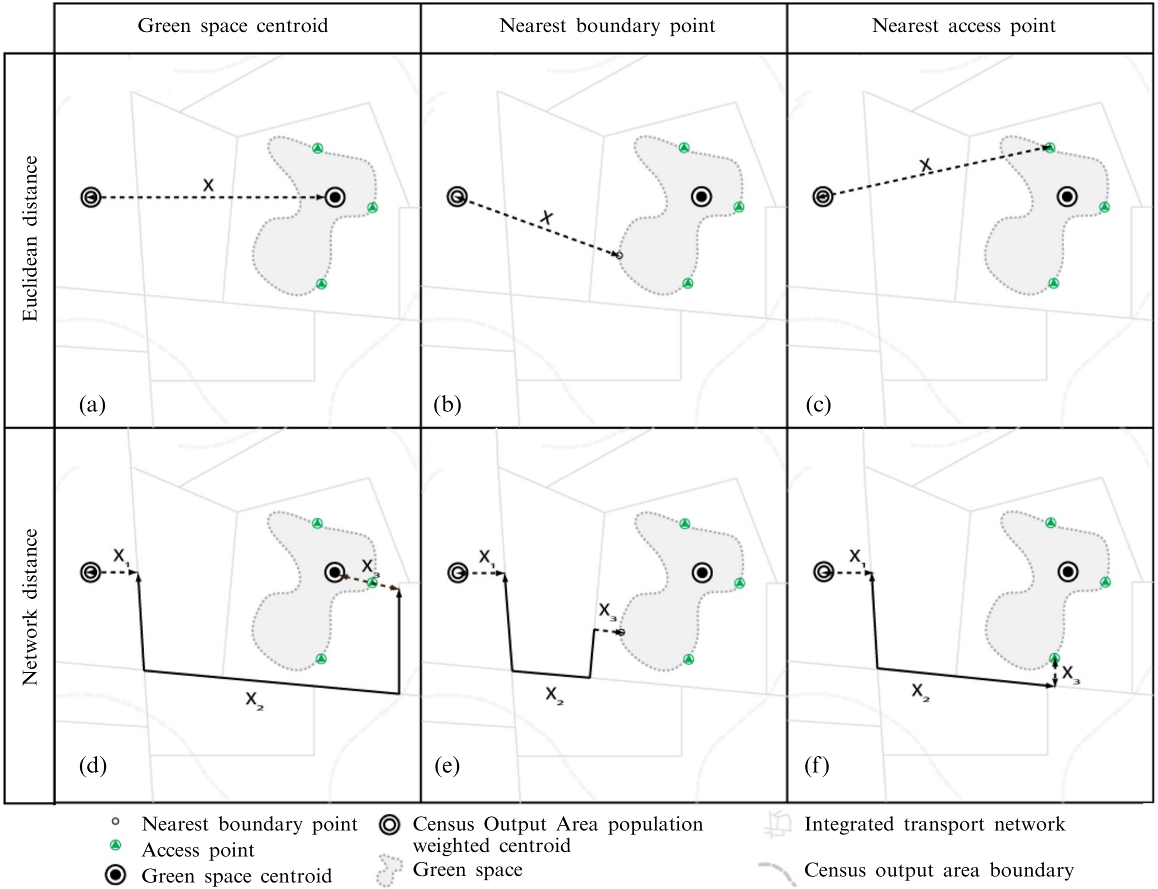
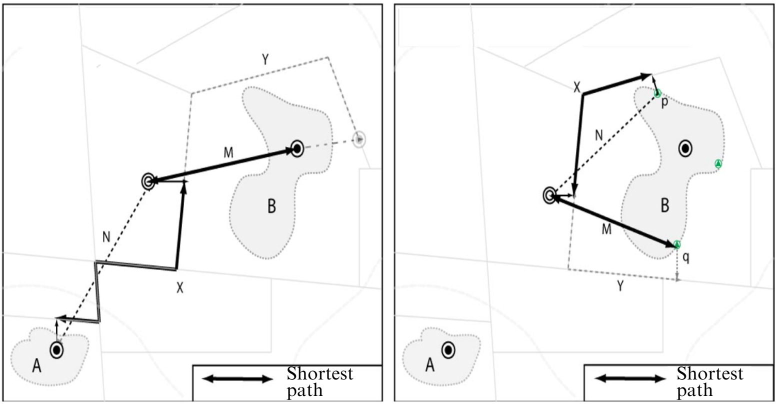


Figure 2. Different approaches to measuring distance to green space. (a) – (f) show scenarios 1 – 6. Descriptions of the scenarios are given in subsection 3.2.



(a) (b)

Figure 3. Complexities in calculating the shortest distance to green space for (a) different green space centroid; (b) same green space, different access point. (See subsections 3.2 for details of lettering.)

closest according to network path distances (ie, $X < Y$). The same effect can arise due to the movement of green space access points when they are adopted as alternative target destinations. The second schematic [figure 3(b)] illustrates how, even when the most proximal green space remains the same, it is possible for a different destination point to become selected as the chosen distance metric is altered. Access point p is selected when using Euclidean distance (since $N < M$), but access point q is selected when using network distance (since $Y < X$). Figure 4 shows an empirical example drawn from our dataset and further illustrates some of the issues involved in determining the nearest green space, and its measured distance, to a given OA. Green space A is selected as being most proximal when based on Euclidean distance measures to either a centroid or the nearest boundary point, while green space B is selected based on the Euclidean distance to a green space access point. Meanwhile green space C has the shortest reported network distance to a centroid, green space A is again associated with the shortest network distance to a boundary point, and finally green space D reports the shortest network distance to an access point. Recording a nearest network distance to a green space boundary is particularly challenging since, as Apparicio and Seguin (2006) note, this requires the identification of a suitable destination point. Their solution was to insert multiple points at a predetermined distance along the entire perimeter, compute the network distance to all such points, and select the one returning the shortest pathway. This solution requires considerable calculation, and it was judged to be computationally intractable when faced with the situation of almost 600 green space polygons, many of which have lengthy perimeters. Instead we elected to use a simpler, if compromised, solution: compute the nearest Euclidean point first and then determine the distance to this same point using both Euclidean and network-based methods. Whilst this ensures the same green space is always identified by either of these approaches, the magnitude of the distance reported to the nearest boundary point can still vary substantially, as shown by the examples in figure 4. A final effect (not illustrated) is that sometimes the 'network path only' distance to a given green space polygon can compute as being shorter than the equivalent Euclidean distance. This may appear counterintuitive but arises as a consequence of the movement of the origin and destination points so that they lie directly on the network dataset.

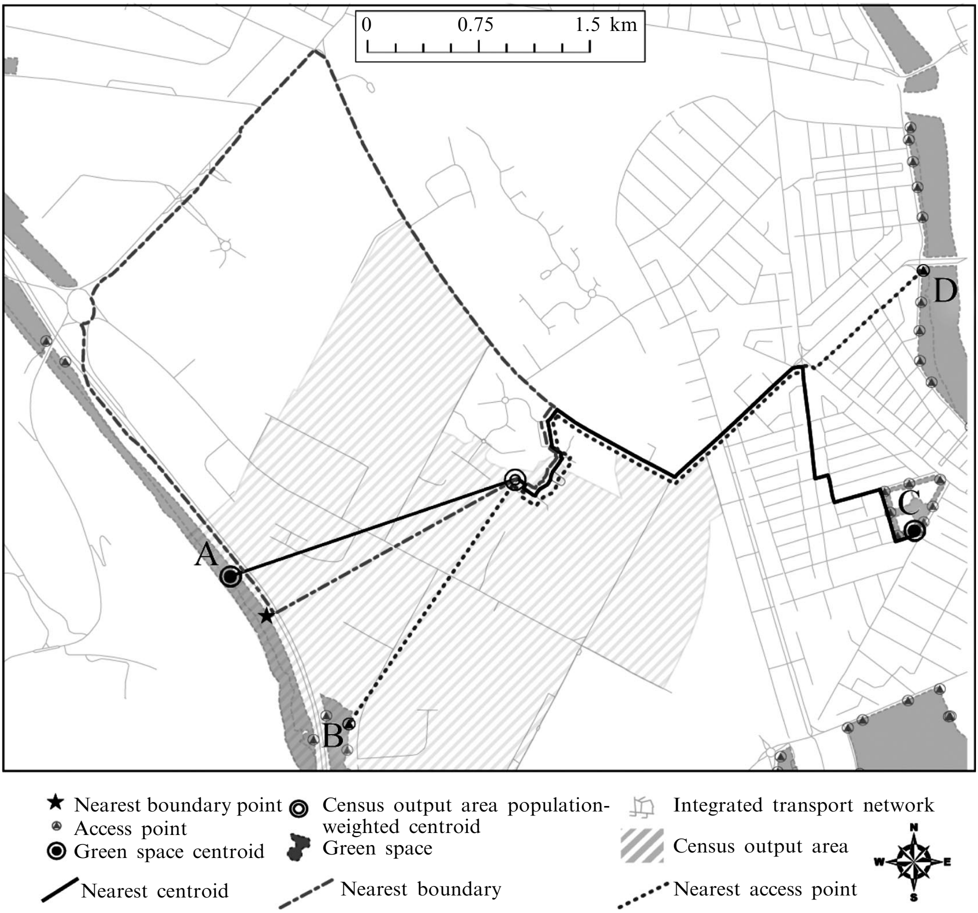


Figure 4. An empirical example of possible 'nearest' green space identifiers and distances.

4 Results

4.1 Comparison of different scenarios

As discussed earlier, GIS-based distance measures are often used to examine associations between green space and health outcomes, or to investigate potential relationships between green space provision and the distribution of socioeconomic, ethnic, and religious groups. Our aim is to explore the implications for any computed levels of association when alternative methods to derive a distance to green space are adopted. In this respect it is the similarity or correlation between respective distance measures that primarily interest us—the more they vary proportionately with each other the less impact that choosing any one method over another is likely to have on any subsequent analyses undertaken. A quick assessment of the statistical distribution of the computed distances revealed all to be highly positively skewed, so we elected to use Spearman's rank correlation coefficient to test for the strength of relationship or dependence.

In table 1 the correlation between Euclidean distances measures to alternative green space target destinations, and the proportion of green space identifiers that overlap, are presented. It can be seen that even when the same distance metric is adopted, the choice of green space target destination point can produce noticeable differences in the pattern of distances that are subsequently calculated, with the lowest correlation ($\rho = 0.835$) arising between geometric centroids and access points

Table 1. Correlations and green space identifier overlap for alternative Euclidean distance measures. Above the diagonal: Spearman's rank correlation coefficient (all significant at $p < 0.001$); below the diagonal: degree (%) to which the same green space object is selected for each output area.

	Green space centroid	Nearest boundary point	Nearest access point
Green space centroid		0.850	0.835
Nearest boundary point	67.3		0.924
Nearest access point	64.8	79.3	

(probably the two most widely used scenarios). Perhaps even more surprising is the extent to which different green space polygons are being selected as being the nearest. At worst more than one third (and at best still more than one fifth) of the selected green spaces do not match for each OA demand population point as alternative destination target points are specified. Once again the biggest difference that arose was between the use of green space geometric centroids and access points. Table 2 presents the equivalent results using network distance measures. Both the strength of the correlations and the proportion of overlapping green space identifiers are seen to diminish even further. The correlation of distances between using a centroid and using the nearest boundary point as the target destination is the weakest, but almost half the green space identifiers do not match when comparing geometric centroids with access points. Furthermore, correlation coefficients were always less than $\rho = 0.80$, and were always lower for full network path distances than for network path only.

Table 2. Correlations and green space identifier overlap for alternative network distance measures. Above the diagonal: Spearman's rank correlation coefficient (all significant at $p < 0.001$); below the diagonal: degree (%) to which the same green space object is selected for each output area.

	Green space centroid	Nearest boundary point	Nearest access point
Green space centroid		0.656[a] 0.608[b]	0.729[a] 0.667[b]
Nearest boundary point	44.9		0.799[a] 0.797[b]
Nearest access point	50.4	60.0	

[a] Network path only.
[b] Full network path.

Tables 3 and 4 examine the differences between Euclidean-based and network-based distances. The closest match ($\rho = 0.853$) arose between the Euclidean and network distance to an access point, whilst the weakest relationship ($\rho = 0.603$) was between Euclidean distance to an access point and full network path distance to a centroid. The degree to which selected green spaces match between Euclidean and network measures often falls below 50%. Values in this table lying along the diagonal whereby the same green space target point is specified are perhaps the most interesting. When using green space centroids only 51% of selected polygons match, whilst when using access points this rises to 70%. The 100% match for the nearest boundary point is a consequence of our experimental design; as explained earlier, to limit the computational load we constrained the Euclidean and network distance calculations to

Table 3. Correlations: Euclidean distance versus network distance: network path only (upper values); full network path (lower values). Spearman's rank correlation coefficient all significant at $p < 0.001$.

Network distance	Euclidean distance		
	green space centroid	nearest boundary point	nearest access point
Green space centroid	0.729	0.665	0.655
	0.665	0.611	0.603
Nearest boundary point	0.623	0.764	0.697
	0.610	0.755	0.698
Nearest access point	0.720	0.804	0.851
	0.714	0.790	0.853

Table 4. Green space identifier overlap: Euclidean distance versus network distance shown as percentages.

Network distance	Euclidean distance		
	green space centroid	nearest boundary point	nearest access point
Green space centroid	50.9	45.1	45.3
Nearest boundary point	67.0	100.0	79.6
Nearest access point	49.6	59.6	70.0

the same 'nearest boundary point'. Whilst we recognised this imparts a bias into our reported results, it is typical of the dilemmas faced by researchers when confronted with the challenge of balancing accuracy and realism with computational practicality. The alternative approach of calculating a network distance to multiple points placed regularly along the boundaries would massively increase the computational workload, would need further justification regarding the interval distance chosen, and is unlikely to be widely contemplated by researchers in such studies (although it may be worthy of further study, especially for small areas). Finally, in table 5 we evaluate the correlations between 'full network path' and 'network path only' distances which suggests that the lowest overall figure ($\rho = 0.604$) arose between the full network path to a centroid and the network path only to the nearest boundary point.

Table 5. Correlations: full network path versus network path only.

Network path only distance	Full network path distance		
	green space centroid	nearest boundary point	nearest access point
Green space centroid	0.911	0.651	0.719
Nearest boundary point	0.604	0.990	0.786
Nearest access point	0.666	0.795	0.984

4.2 Possible association between scenarios and a measure of deprivation

Since the focus of this study is on how competing methodologies for determining green space proximity might impact upon computed levels of association with other confounding variables, these correlation statistics (and proportions of overlapping identifiers) provide the best general evidence that choice of distance metric and green space representation are nontrivial and should at least be carefully reported. The impacts of such choices on any subsequent analyses will always be case specific, but the results presented in table 6 help to illustrate their potential importance. Table 6 shows Spearman's rank correlation coefficients between selected green space distance metrics and a census-based index of material deprivation which is often used in studies relating health data with socioeconomic conditions, the Townsend deprivation index [see Senior (2002) for more details]. Although strength of association in this example is low in all cases, it varies from a high of $\rho = 0.193$ to a low of $\rho = 0.082$ depending on the specific distance metric and green space representation adopted. It is possible to conceive how, when applied in other studies, such variability might make all the difference in obtaining statistically significant outcomes.

Table 6. Correlations between the Townsend deprivation index (2001 output areas) and distance metrics versus target destinations.

Distance metric	Target destination		
	green space centroid	nearest boundary point	nearest access point
Euclidean distance	0.186	0.186	0.169
Network path only distance	0.136	0.108	0.193
Full network path distance	0.082	0.129	0.185

As a further example, using the OA Townsend scores we divided the study area's census tracts into deprivation quintiles. The distance to the nearest green space between the uppermost and lowermost groups could then be compared. We postulate that people living in the most deprived census tracts will have to travel further to access a green space than those living in the least deprived census tracts. Figure 5 illustrates the profile of average distance scores across the quintile groups using Euclidean and 'full network path' distances. It appears that our hypothesis is supported in the case of a Euclidean distance metric (the average distance for the most deprived tracts is 328 m, versus 238 m for the least deprived), but the evidence is much less convincing for 'full network path' distances (587 m versus 555 m). The results from a one-tailed independent t-test, in which equal variances were not assumed, and with

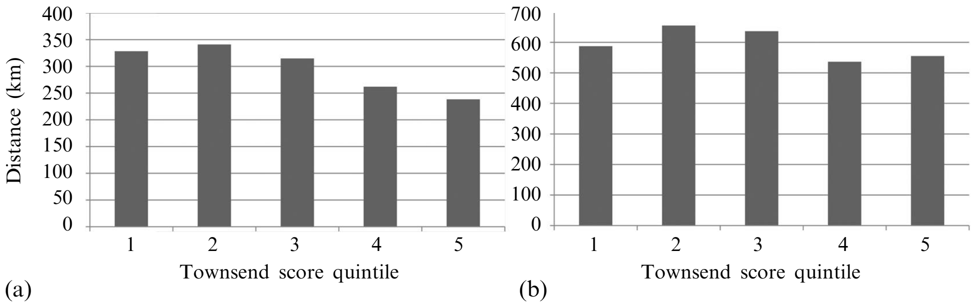


Figure 5. Mean distance to nearest green space, categorised by Townsend score quintiles; (a) Euclidean distance; (b) full network path distance.

distance replaced by square-root distance to counteract the positive skew of the data values, further supports this outcome. The Euclidean distance to centroid metric reports a significant difference in mean scores ($t = 5.07$, $p < 0.01$) between the most and least deprived census tracts, whilst the full network path metric returns a result of no significant difference ($t = 0.072$, $p > 0.05$). This in turn suggests the way in which distance is computed in such studies can have an influence on any potential relationships with socioeconomic conditions (as measured by the Townsend index) and thus it should be 'best practice' for researchers exploring such differences to incorporate a range of methodological approaches to examine the robustness of any associations.

5 Discussion

An increasing number of studies are concerned with measuring the availability of, and access to, green spaces from either environmental justice, inequalities in provision or health outcomes perspectives. For example, access to green space has been associated with less stress and lower body mass index amongst adults (Nielsen and Hansen, 2007), improved self-reported health (Mitchell and Popham, 2007), and shorter postoperative recovery periods (Ulrich, 1984). Amongst children and youths, the positive health effects of green landscapes include a significant increase in moderate-to-vigorous physical activity. However, few studies have explored the implications of using alternative measures of accessibility based either on network or Euclidean distance and of different representations of green space (eg centroid, access points, and boundary points) in the analysis of potential relationships with health variables or socioeconomic measures. This paper has considered the sensitivity of such accessibility measures to the ways in which distance is estimated. Our findings, based on a 'typical' case study, suggest that alternative approaches result in differences in distance calculations which may be important in determining the strength of relationship with material conditions as measured by a census-based deprivation measure. Of immediate importance for those engaged in researching potential links between access to green spaces and health outcomes, the study highlights that in a significant number of instances different green spaces are identified as being 'the closest' depending on the techniques used within GIS-based methodologies. The use of different measurement techniques suggests that researchers may well not be matching on the same green spaces, which we suggest may be important where relationships are sought with the attributes of the nearest available green space, which in turn may influence the direction (and strength) of any association.

Whilst we would concur with Sander et al (2010) that network-based potential accessibility measures improve on a basic Euclidean analysis we recognise that in health contexts such measures only approximate to actual levels of exposure to green spaces. In reality different sets of factors may be influencing the use of green spaces and may in certain situations lead to groups accessing more desirable parks at further distances (Veitch et al, 2008). For example, recent studies have drawn attention to the potential role of neighbourhood social characteristics, in particular factors such as perceived crime and traffic safety, in influencing the extent to which facilities are deemed accessible, particularly to children (eg, Loukaitou-Sideris and Sideris, 2010), and the elderly (Sugiyama et al, 2009). Jones et al (2009) have compared perceived access in relation to green space provision in Bristol and found a discrepancy for residents in more deprived areas, suggesting distance is not the only consideration in predicting usage levels by those residents. Relationships between perceptions of neighbourhood environment, distance to services and facilities, and physical activity levels and ultimately health outcomes continue to form the focus of a number of studies (eg, Panter and Jones, 2008).

No attempt has been made in the current study to estimate accessibility for particular population subsets (eg, children), to analyse the data by their demographic characteristics or for different ethnic groups (eg, Comber et al, 2008), or restrict analysis by the type of green space. However, these are clearly areas for further research given concerns over, for example, levels of childhood obesity and inequity of provision amongst different social groups. McCormack et al (2010) call for the integration of qualitative methods with quantitative findings to understand more fully the factors influencing urban park use, while Wridt (2010) investigated qualitative GIS approaches to examining children's access to play areas in Denver, USA. A final research area that follows on from the analysis conducted here relates to the quality of such spaces. Factors such as the quality of individual parks, opening times, and neighbourhood effects could be incorporated into wider studies of the relationships between health and green space within GIS-based analysis. Findings from our research have drawn attention to the need to consider a range of different accessibility techniques in providing objective measures of access within such studies. The impacts of changes over time in the characteristics of green spaces and how this may influence temporal variations in accessibility could form the basis for further research in this area (see for example, Kessel et al, 2009).

6 Conclusions

Whilst it is recognised that dimensions of access to green spaces, as to other services and facilities, is multifaceted and by no means just confined to geographical factors, some of the previous studies reviewed in this paper present evidence of spatial variations in, for example, physical activity levels in relation to potential access to such opportunities. We contend that, whilst there is an increasing literature concerning areas such as environmental justice and health inequalities that is focusing on the relationship between environmental factors and a wide variety of quality-of-life measures including health outcomes, few studies have highlighted the implications of using alternative measures of accessibility on the degree of association between access to green spaces and such variables. Intuition would suggest that, in the absence of detailed utilisation patterns on the actual spaces being used, the 'gold standard' when measuring potential accessibility to the nearest green space using proxy measures would involve measuring access from individual household locations to public entrances (or other access points) to green spaces using network distance (using as detailed a path or road network as is available). However, a review of the existing literature highlights a lack of consistency in the approaches taken to measure accessibility to green spaces which we suggest may impact on the potential associations found with socioeconomic or health outcomes. Few studies incorporate a range of accessibility measures based on alternative methodologies when examining such relationships and many ignore the consequences of the ways in which such calculations are performed using standard algorithms and methods of spatial representation within GIS models.

In this study we illustrate such impacts using a database of green spaces and access points for a 'typical' urban area in the UK to demonstrate the implications of underplaying the importance of such factors. Using real-world examples we suggest that these may have dramatic impacts on seemingly straightforward calculations involving for example measuring access to the 'nearest' green space. This is demonstrated by firstly examining levels of association between different distance measures based on alternative methodologies using correlation statistics, and secondly by using similar techniques to examine variations in the relationship with an index of material deprivation that is widely used in studies of health variations in the UK. Our findings suggest that any analysis of the relationships between access measures, health variables, and the attributes of such green spaces may be fundamentally flawed unless the consequences of alternative methodological approaches are at least highlighted and sensitivity analyses conducted. In our future research we aim to develop alternative measures of accessibility to green space by incorporating a number of methodological advancements using, for example, detailed population estimation models to measure access to green spaces. This in turn is part of a wider project aimed at measuring access to a more extensive range of community resources for different sections of the population in order to investigate potential associations with health outcomes. In this research we suggest that the methods by which distance is calculated may influence the strength of the relationship with a standard measure of material deprivation. In our future research we will investigate whether there are any variations in relationships between alternative measures of distance and different health variables. In the absence of data on the detailed spatial and temporal utilisation patterns of green spaces, we have been unable to compare the potential measures developed here with actual use of green spaces in Cardiff. However, our research has cautioned against the use of inappropriate methodologies in examining access to green space, which may directly influence directions (and levels) of association and hence may limit their relevance in wider geographical contexts.

Acknowledgements. This paper is based on research supported by the Wales Institute of Social and Economic Research, Data and Methods (WISERD), funded by the Economic and Social Research Council (ESRC) (Grant reference: RES-576-25-0021) and the Higher Education Funding Council for Wales (HEFCW).

References

Abercrombie L, Sallis J, Terry L, Lawrence F, Saelens B, Chapman J, 2008, "Income and racial disparities in access to public parks and private recreation facilities" *Preventative Medicine* **34** 9 – 15

Apparicio P, Seguin A, 2006, "Measuring the accessibility of services and facilities for residents of public housing in Montreal" *Urban Studies* **43** 187 – 211

Apparicio P, Abdelmajid M, Riva M, Shearmur R, 2008, "Comparing alternative approaches to measuring the geographical accessibility of urban health services: distance types and aggregation error issues" *International Journal of Health Geographics* **7**:7, doi:10.1186/1476-072X-7-7

Barbosa O, Tratalos J, Armsworth P, Davies R, Fullers R, Johnson P, Gaston K, 2007, "Who benefits from access to green space? A case study from Sheffield, UK" *Landscape and Urban Planning* **83** 187 – 195

Boone C G, Buckley G L, Grove J M, Sister C, 2009, "Parks and people: an environmental justice inquiry in Baltimore, Maryland" *Annals of the Association of American Geographers* **99** 767 – 787

Cardiff County Council, 2008 *An analysis of Accessible Natural Green Space Provision in Cardiff* Policy Group, Strategic Planning and Environment, Cardiff County Council

Comber A J, Brunsdon C, Green E, 2008, "Using a GIS-based network analysis to determine urban greenspace accessibility for different ethnic and religious groups" *Landscape and Urban Planning* **86** 103 – 114

Coombes E, Jones A, Hillsdon M, 2010, "The relationship of physical activity and overweight to objectively measured green space accessibility and use" *Social Science and Medicine* **70** 816 – 822

de Vries S, Verheij R, Groenewegen P, Spreeuwenberg P, 2003, "Natural environments—healthy environments? An exploratory analysis of the relationship between green space and health" *Environment and Planning A* **35** 1717 – 1731

Forsyth A, 2000, "Analysing public space at a metropolitan scale: notes on the potential for using GIS" *Urban Geography* **21** 121 – 147

Giles-Corti B, Broomhall M H, Knuiman M, Collins C, Douglas K, Ng K, Lange A, Donovan R J, 2005, "Increasing walking: how important is distance to, attractiveness and size of public open space?" *American Journal of Preventative Medicine* **28** 169 – 176

Hewko J, Smoyer-Tomic K, Hodgson M J, 2002, "Measuring neighbourhood spatial accessibility: does aggregation error matter?" *Environment and Planning A* **34** 1185 – 1206

Humpel N, Owen N, Leslie E, 2002, "Environmental factors associated with adults' participation in physical activity: a review" *American Journal of Preventative Medicine* **22** 188 – 199
Jones A, Hillsdon M, Coombes E, 2009, "Greenspace access, use and physical activity: understanding the effects of area deprivation" *Preventive Medicine* **49** 500 – 505
Jones S G, Ashby A J, Momin S R, Naidoo A, 2010, "Spatial implications associated with using Euclidean distance measurements and geographic centroid imputation in health care research" *Health Services Research* **45** 316 – 327
Jordan H P, Roderick P, Martin D, Barnett S, 2004, "Distance, rurality and the need for care: access to health services in South West England" *International Journal of Health Geographics* **3** 21, doi:10.1186/1476-072X-3-21
Kaczynski A, Henderson K A, 2007, "Environmental correlates of physical activity: a review of evidence about parks and recreation" *Leisure Sciences* **29** 315 – 354
Kaczynski A, Potwarka L R, Saelens B, 2008, "Association of park size, distance, and features with physical activity in neighbourhood parks" *American Journal of Public Health* **98** 1451 – 1456
Kessel A, Green J, Pinder R, Wilkinson P, Grundy C, Lachowycz K, 2009, "Multidisciplinary research in public health: a case study of research on access to green space" *Public Health* **123** 32 – 38
Loukaitou-Sideris A, Sideris A, 2010, "What brings children to the park? Analysis and measurement of the variables affecting children's use of parks" *Journal of the American Planning Association* **76** 89 – 107
Maas J, Verheij R, de Vries S, Spreeuwenberg P, Schellevis F G, Groenewegen P P, 2009, "Morbidity is related to a green living environment" *Journal of Epidemiology and Community Health* **63** 967 – 973
McCormack G R, Rock M, Toohey A M, Hignell D, 2010, "Characteristics of urban parks associated with park use and physical activity: a review of qualitative research" *Health and Place* **16** 712 – 726
Macintyre S, MacDonald L, Ellaway A, 2008, "Do poorer people have poorer access to local resources and facilities? The distribution of local resources by area deprivation in Glasgow, Scotland" *Social Science and Medicine* **67** 900 – 914
Mitchell R, Popham F, 2007, "Green space, urbanity and health: relationships in England" *Journal of Epidemiology and Community Health* **61** 681 – 683
Mitchell R, Popham F, 2008, "Effect of exposure to natural environment on health inequalities: an observational population study" *The Lancet* **372** 1655 – 1660
Nicholls S, 2001, "Measuring the accessibility and equity of public parks: a case study using GIS" *Managing Leisure* **6** 201 – 219
Nicholls S, Shafer S, 2001, "Measuring accessibility and equity in a local park system: the utility of geospatial technologies to park and recreation professionals" *Journal of Park and Recreation Administration* **19** 102 – 124
Nielsen T S, Hansen K B, 2007, "Do green areas affect health? Results from a Danish survey on the use of green areas and health indicators" *Health and Place* **13** 839 – 850
Panter J R, Jones A P, 2008, "Associations between physical activity, perceptions of the neighbourhood environment and access to facilities in an English city" *Social Science and Medicine* **67** 1917 – 1923
Phibbs C S, Luft H S, 1995, "Correlation of travel time on roads versus straight line distance" *Medical Care Research and Review* **52** 532 – 542
Powell L, Slater S, Chaloupka F, 2004, "The relationship between community physical activity settings and race, ethnicity and socioeconomic status" *Evidence-based Preventative Medicine* **1** 135 – 144
Sander H A, Ghosh D, van Riper D, Manson S M, 2010, "How do you measure distance in spatial models? An example using open-space valuation" *Environment and Planning B: Planning and Design* **37** 874 – 894
Senior M, 2002, "Deprivation indicators", in *The Census Data System* Eds P Rees, D Martin, P Williamson (John Wiley, Chichester, Sussex) pp 123 – 137
Sugiyama T, Ward-Thompson C, Alves S, 2009, "Associations between neighbourhood open space attributes and quality of life for older people in Britain" *Environment and Behavior* **41** 3 – 21
Talen E, Anselin L, 1998, "Assessing spatial equity; an evaluation of measures of accessibility to public playgrounds" *Environment and Planning A* **30** 595 – 613
Ulrich RS, 1984, "View through a window may influence recovery from surgery" *Science* **224** (4647) 420 – 421

Van Herzele A, Wiedemann T, 2003, "A monitoring tool for the provision of accessibility and attractive urban green spaces" *Landscape and Urban Planning* **63** 109 – 126
Veitch J, Salmon J, Ball K, 2008, "Children's active free-play in local neighbourhoods: a behavioural mapping study" *Health Education Research* **23** 870 – 879
Wridt P, 2010, "A qualitative GIS approach to mapping urban neighborhoods with children to promote physical activity and child-friendly community planning" *Environment and Planning B: Planning and Design* **37** 129 – 147